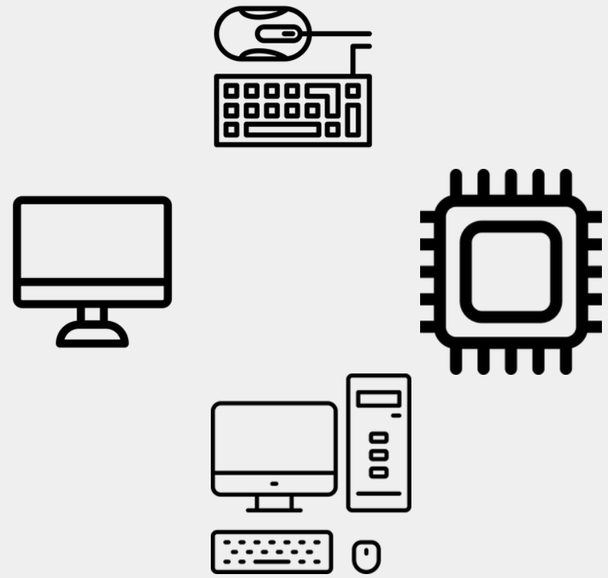
Data Analytics Project using Sql



CONSUMER GOODS AD_HOC INSIGHTS FOR ATLIQ HARDWARE

Created by Keyur Dattani

Agenda

- About company
- Problem Statement and Objectives
- Data received from company
- 10 ad hoc requests received from the company and generated outputs and insights

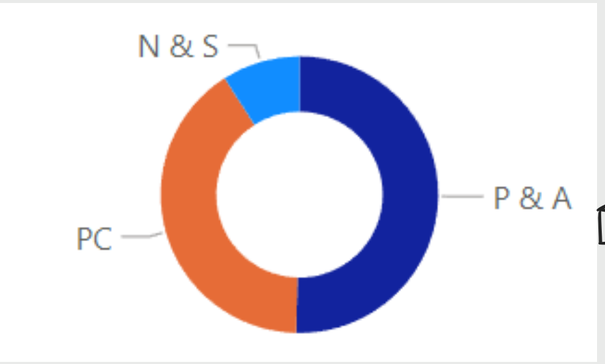
About company:

Atliq Hardware, headquartered in India, stands out as a significant player in the computer hardware industry. Its influential presence extends beyond the borders of India, encompassing numerous other nations.

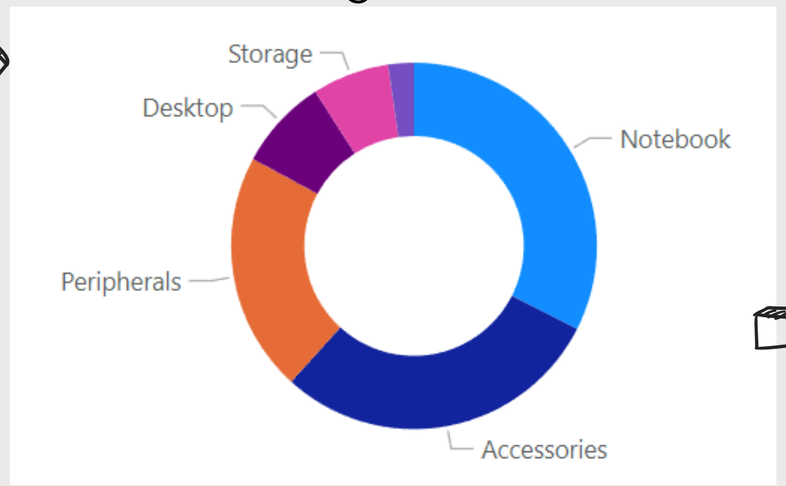
Markets of Atliq Hardware

Categories and types of product sold by Atliq hardware

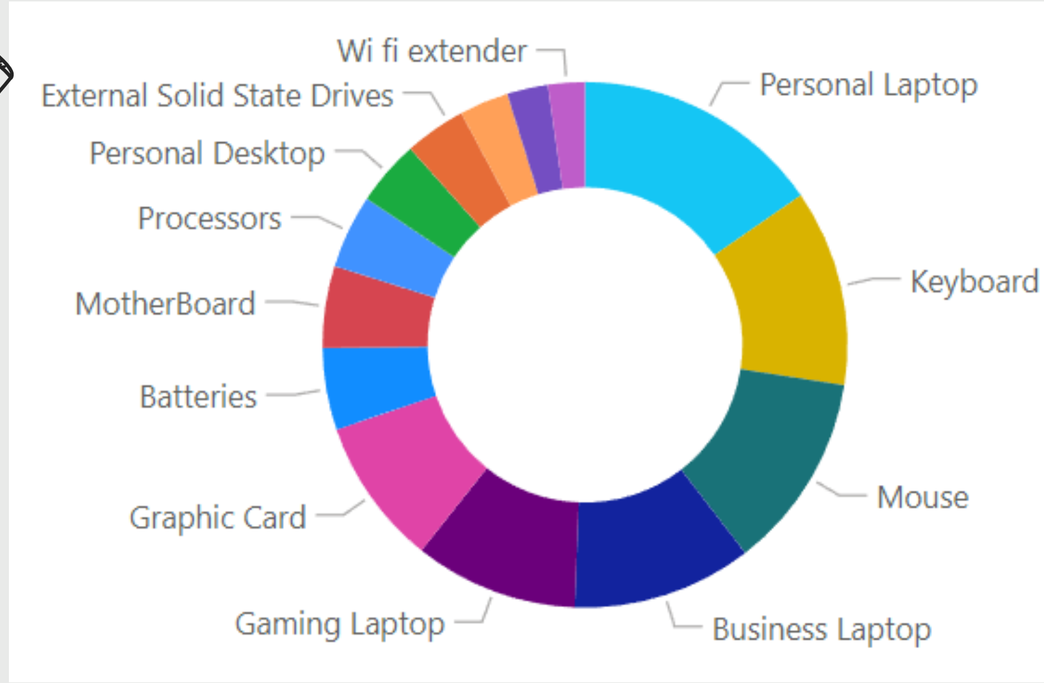
Division



Segment



Categories



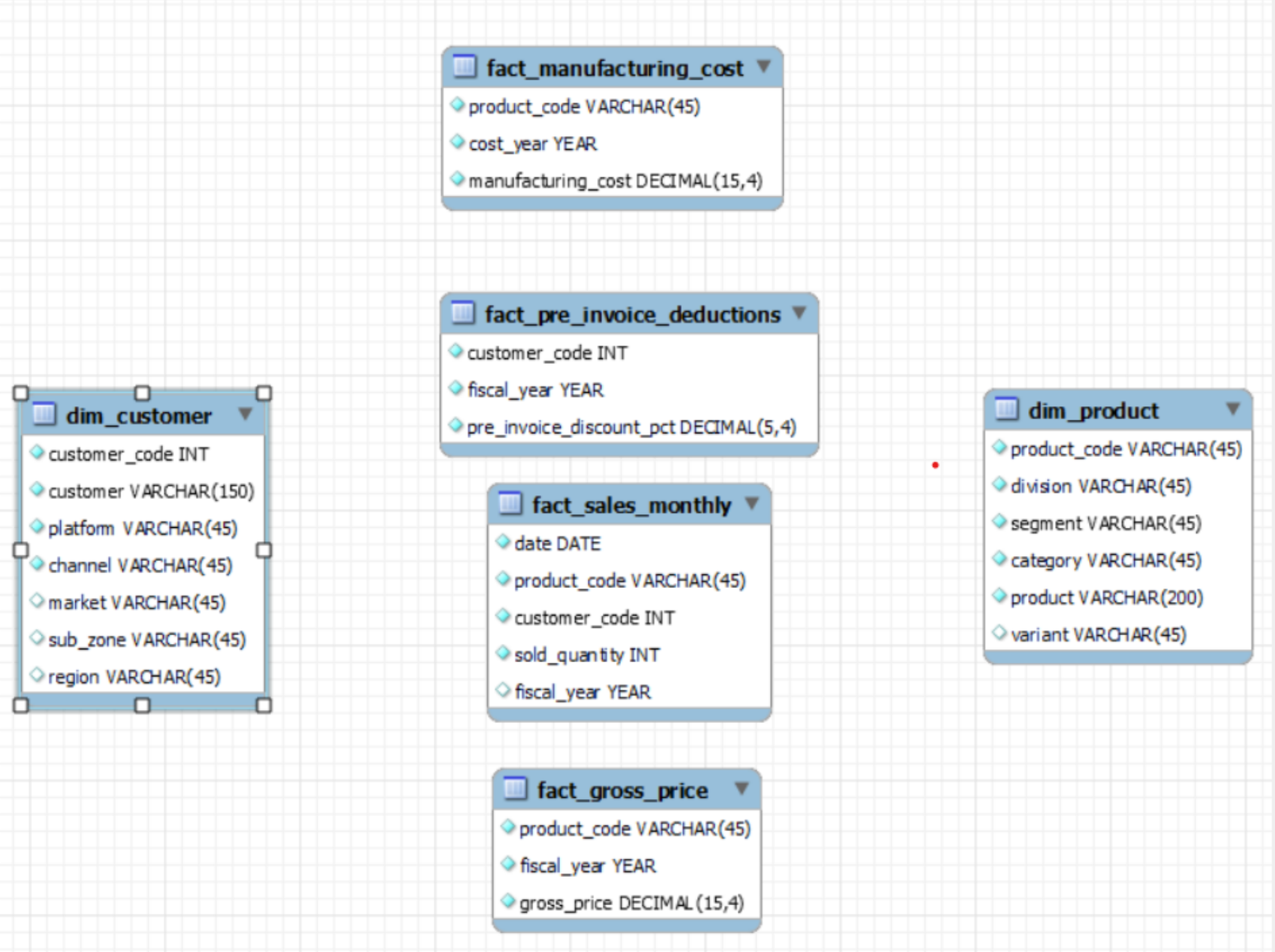
Problem Statement and Objectives:

Atliq Hardware: Leading computer hardware manufacturer in India is expanding globally and the management acknowledges lack of timely data-driven decisions.

- Initiative: Strengthen data analytics team with junior analysts.
- Director Tony Sharma's approach: Conduct SQL challenge for tech and soft skills assessment.
- Objective: Address 10 ad hoc data requests to gain valuable insights
- Focus: Enhance decision-making through data-driven strategies.

Data received from the company:

Input data consists of sales data for FY 2020 and FY 2021, along with different other dimension tables like customer details, product details, etc.



Ad_hoc requests received :

Codebasics SQL Challenge

Requests:

1. Provide the list of markets in which customer "Atliq Exclusive" operates its business in the APAC region.

2. What is the percentage of unique product increase in 2021 vs. 2020? The final output contains these fields,
   - unique_products_2020
   - unique_products_2021
   - percentage_chg

3. Provide a report with all the unique product counts for each segment and sort them in descending order of product counts. The final output contains 2 fields,
   - segment
   - product_count

4. Follow-up: Which segment had the most increase in unique products in 2021 vs 2020? The final output contains these fields,
   - segment
   - product_count_2020
   - product_count_2021
   - difference

5. Get the products that have the highest and lowest manufacturing costs. The final output should contain these fields,
   - product_code
   - product
   - manufacturing_cost

6. Generate a report which contains the top 5 customers who received an average high pre_invoice_discount_pct for the fiscal year 2021 and in the Indian market. The final output contains these fields,
   - customer_code
   - customer
   - average_discount_percentage

7. Get the complete report of the Gross sales amount for the customer **"Atliq Exclusive"** for each month. This analysis helps to get an idea of low and high-performing months and take strategic decisions.
   The final report contains these columns:
   - Month
   - Year
   - Gross sales Amount

8. In which quarter of 2020, got the maximum total_sold_quantity? The final output contains these fields sorted by the total_sold_quantity,
   - Quarter
   - total_sold_quantity

9. Which channel helped to bring more gross sales in the fiscal year 2021 and the percentage of contribution? The final output contains these fields,
   - channel
   - gross_sales_mln
   - percentage

10. Get the Top 3 products in each division that have a high total_sold_quantity in the fiscal_year 2021? The final output contains these fields,
    - division
    - product_code

Tools Used

MySQL

Excel

Power BI

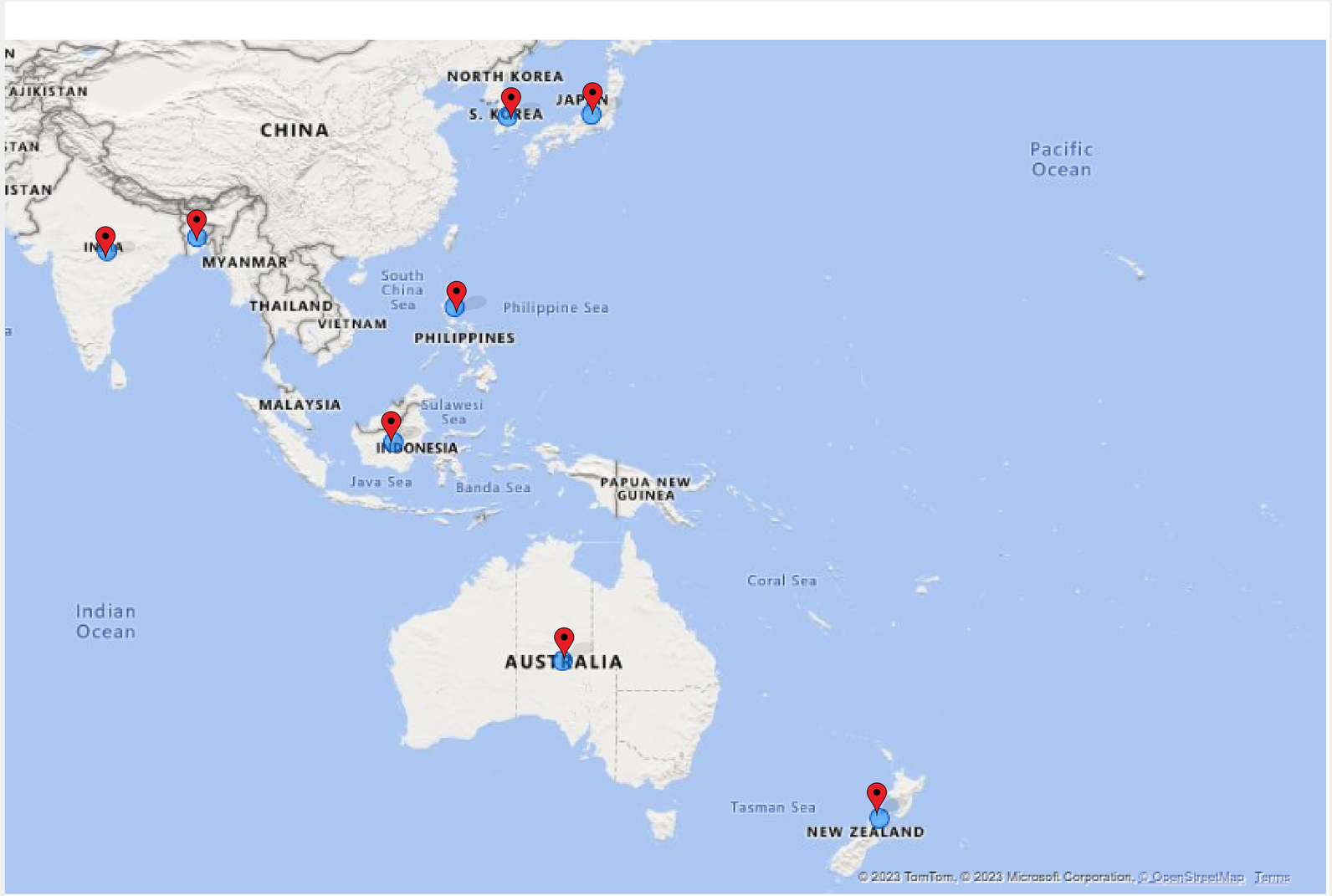
1. Provide the list of markets in which customer "Atliq Exclusive" operates its business in the APAC region.

```sql
SELECT market FROM gdb023.dim_customer
where customer = "Atliq Exclusive" and region="APAC";
```

market
India
Indonesia
Japan
Philiphines
South Korea
Australia
Newzealand
Bangladesh

Insights:

In APAC region Atliq Exclusive opreates its business in 8 major countries they are India ,Indonesia,Japan,Philiphines,South Korea,Australia,New Zealand,Bangladesh



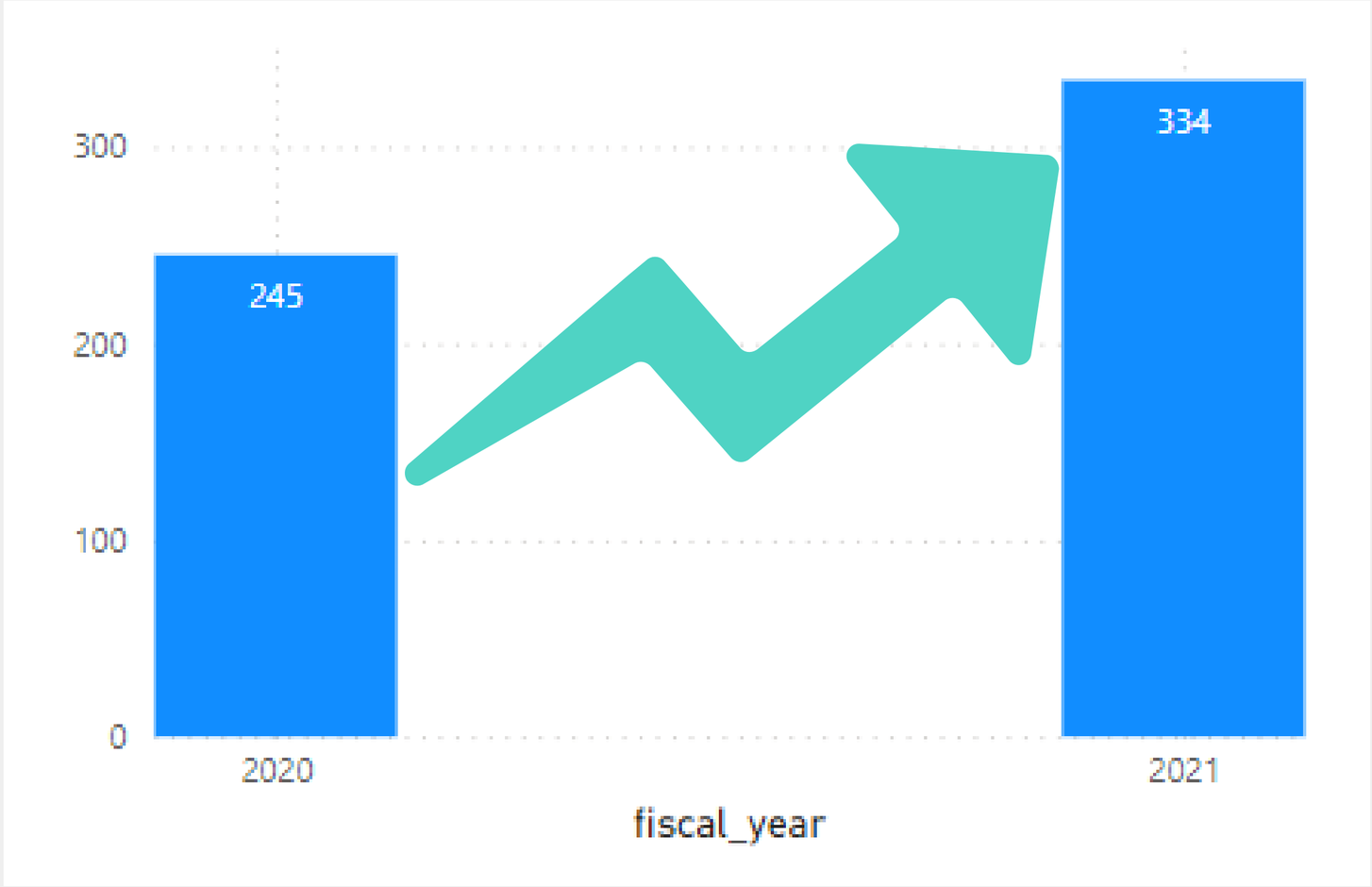
2. What is the percentage of unique product increase in 2021 vs. 2020? The final output contains these fields, unique_products_2020 unique_products_2021 percentage_chg

```sql
WITH ProductCounts AS (
    SELECT
        fiscal_year,
        COUNT(DISTINCT product_code) AS unique_products
    FROM
        gdb023.fact_sales_monthly
    GROUP BY
        fiscal_year
)
SELECT
    p2020.unique_products AS unique_products_2020,
    p2021.unique_products AS unique_products_2021,
    CASE
        WHEN p2020.unique_products = 0 THEN NULL
        ELSE ((p2021.unique_products - p2020.unique_products) / p2020.unique_products) * 100
    END AS percentage_chg
FROM
    ProductCounts p2020
JOIN
    ProductCounts p2021 ON p2020.fiscal_year = 2020 AND p2021.fiscal_year = 2021;
```

	unique_products_2020	unique_products_2021	percentage_chg
▶	245	334	36.3265

Insights:

The rise in unique products gives us insights that the company has strong alignment with market demand and Diversifying product offerings will empower customers with more choices, potentially driving increased sales and enhancing overall customer contentment.



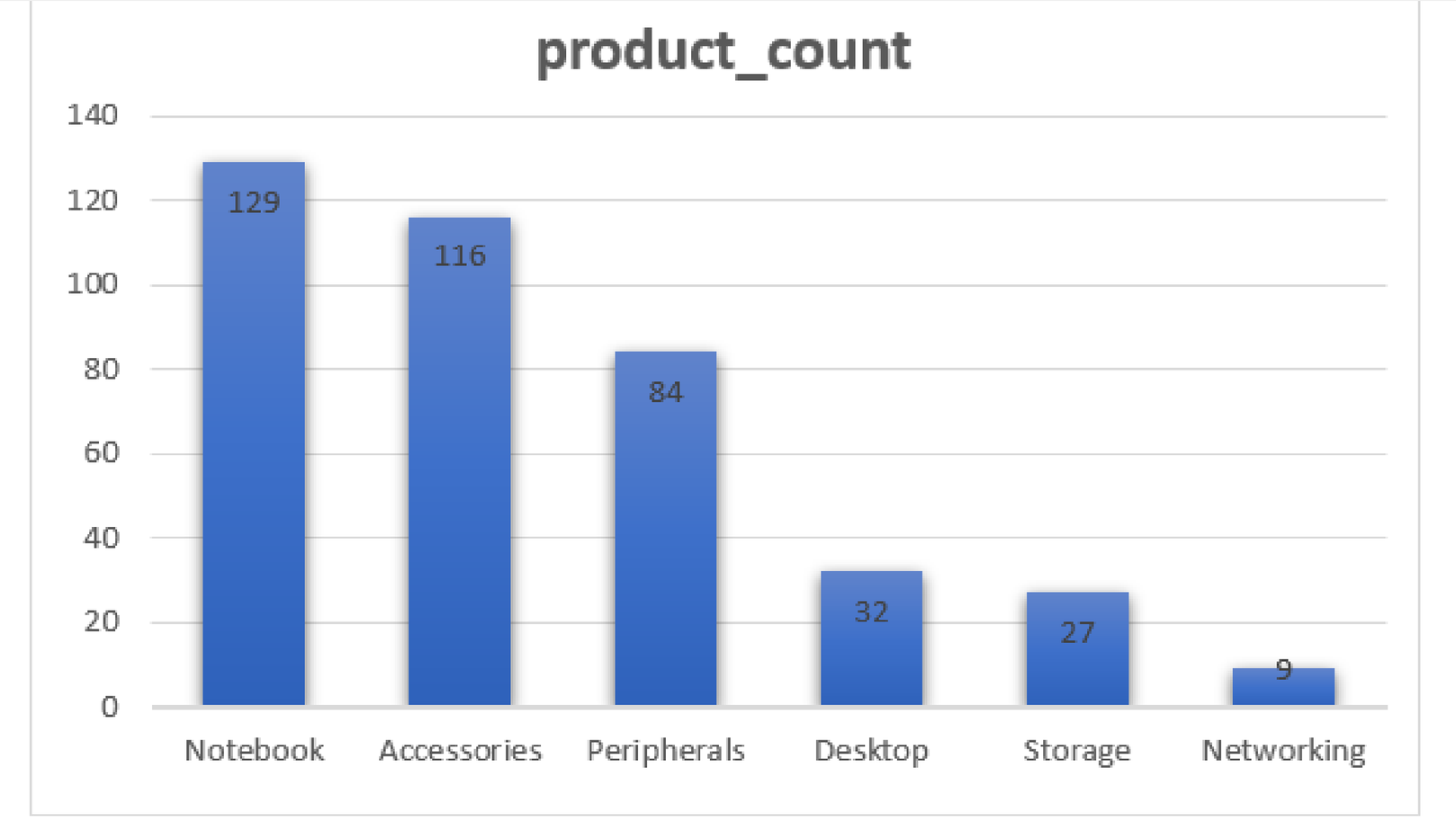
3. Provide a report with all the unique product counts for each segment and sort them in descending order of product counts. The final output contains 2 fields, segment product_count

```
SELECT segment,count(distinct product_code) as product_count FROM gdb023.dim_product
group by segment
order by product_count desc;
```

segment	product_count
Notebook	129
Accessories	116
Peripherals	84
Desktop	32
Storage	27
Networking	9

Insights:

Segments such as notebooks, accessories, and peripherals are showing significant manufacturing growth as compared to desktops, storage, and networking this could be due to greater demand for innovation in notebooks, accessories, and peripherals



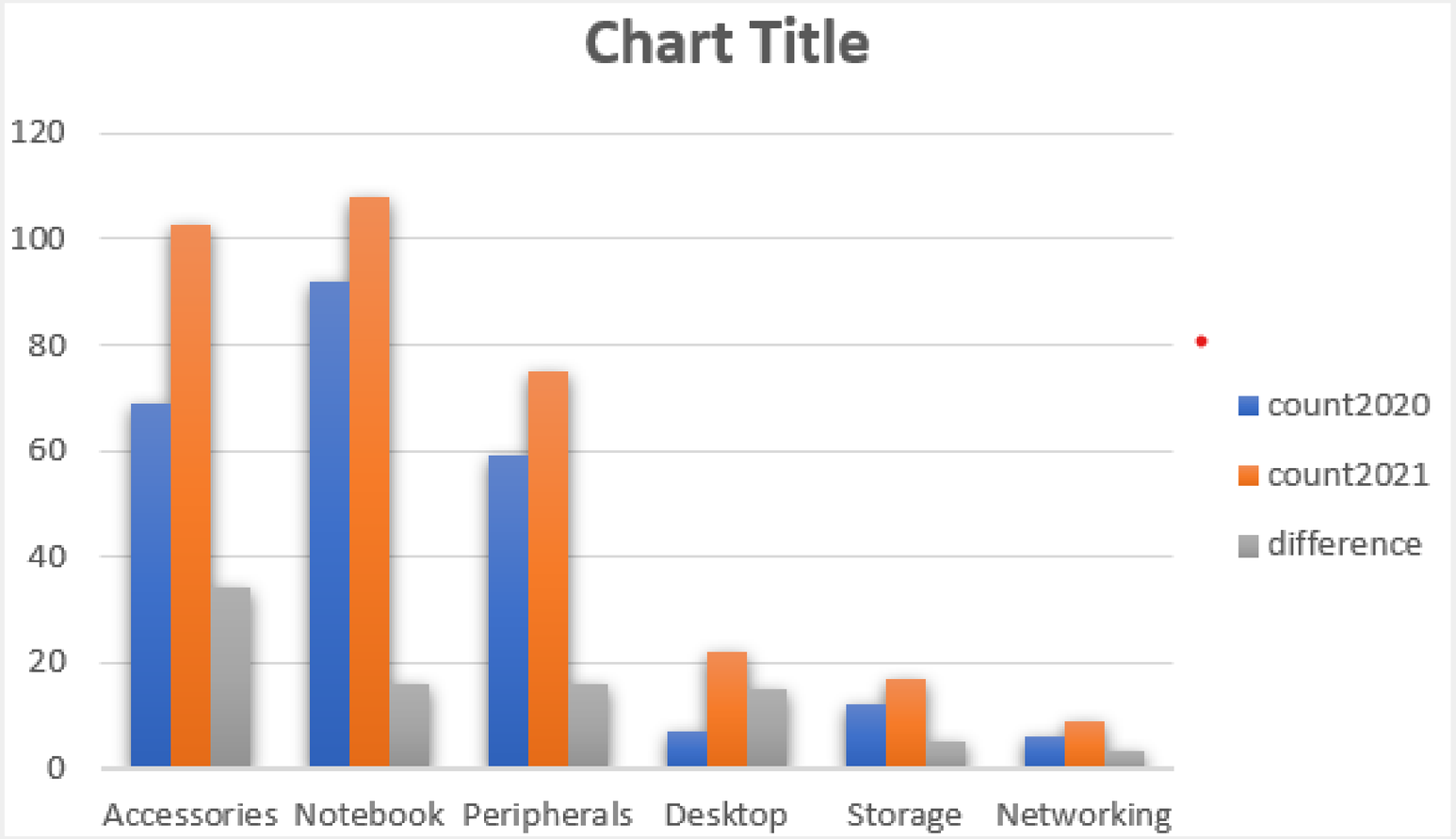
4. Follow-up: Which segment had the most increase in unique products in 2021 vs 2020? The final output contains these fields, segment product_count_2020, product_count_2021, difference .

```sql
with c2020 as (
    SELECT dp.segment, COUNT(DISTINCT fs.product_code) AS count2020
FROM fact_sales_monthly fs
JOIN dim_product dp ON fs.product_code = dp.product_code
WHERE fs.fiscal_year = 2020
GROUP BY dp.segment

),
c2021 as (
   SELECT dp.segment, COUNT(DISTINCT fs.product_code) AS count2021
FROM fact_sales_monthly fs
JOIN dim_product dp ON fs.product_code = dp.product_code
WHERE fs.fiscal_year = 2021
GROUP BY dp.segment

)

select c.segment,c.count2020, d.count2021, (d.count2021 - c.count2020) as difference
from c2020 c
join c2021 d
on c.segment=d.segment
group by c.segment
order by difference desc;
```

segment	count2020	count2021	difference
Accessories	69	103	34
Notebook	92	108	16
Peripherals	59	75	16
Desktop	7	22	15
Storage	12	17	5
Networking	6	9	3

Insights:

- In the year 2021 the company's primary focus was on enhancing the diversity of our accessories segment. During this period the successfully launched 34 new products within the accessories category
- But the lowest production growth was for Networking and storage category



5. Get the products that have the highest and lowest manufacturing costs. The final output should contain these fields, product_code ,product manufacturing_cost

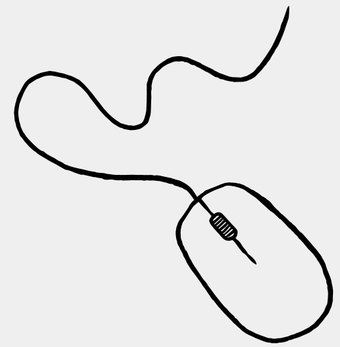
```sql
SELECT p.product_code, p.product, m.manufacturing_cost
FROM dim_product p
JOIN fact_manufacturing_cost m ON p.product_code = m.product_code
where m.manufacturing_cost in (select min(manufacturing_cost)
from fact_manufacturing_cost
union
select max(manufacturing_cost) from fact_manufacturing_cost);
```

product_code	product	manufacturing_cost
A2118150101	AQ Master wired x1 Ms	0.8920
A6120110206	AQ HOME Allin1 Gen 2	240.5364

Insights:

Among our product manufacturing costs, Desktop having product code:A6120110206 is having highest manufacturing cost which is 240.54 dollars,having the highest expenditure, whereas Mouse having product code:A2118150101 is having the lowest production cost.

$240.54
A6120110206
Personal Desktop



$0.89
A2118150101
Mouse

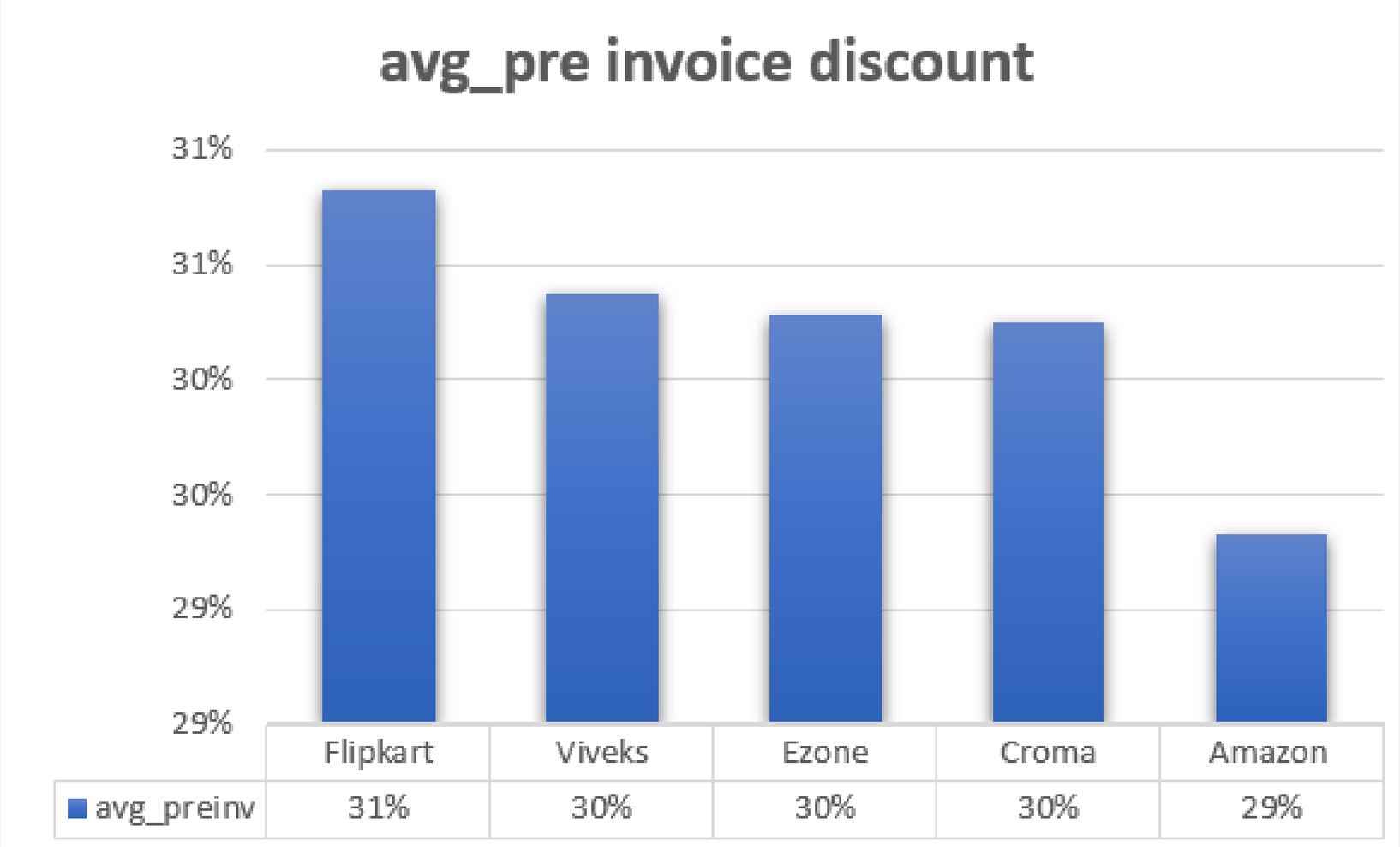
6. Generate a report which contains the top 5 customers who received an average high pre_invoice_discount_pct for the fiscal year 2021 and in the Indian market. The final output contains these fields, customer_code,customer average_discount_percentage

```
SELECT fd.customer_code,dc.customer,avg(fd.pre_invoice_discount_pct)
as avg_preinv
FROM gdb023.fact_pre_invoice_deductions fd
join dim_customer dc on
fd.customer_code=dc.customer_code
where fd.fiscal_year=2021 and dc.market="India"
group by fd.customer_code,dc.customer
order by avg_preinv desc
limit 5;
```

customer_code	customer	avg_preinv
90002009	Flipkart	30.83000000
90002006	Viveks	30.38000000
90002003	Ezone	30.28000000
90002002	Croma	30.25000000
90002016	Amazon	29.33000000

Insights:

- The highest average pre-invoice discount (31%) was given to Flipkart.
- The lowest average pre-invoice discount(29%) was given to Amazon



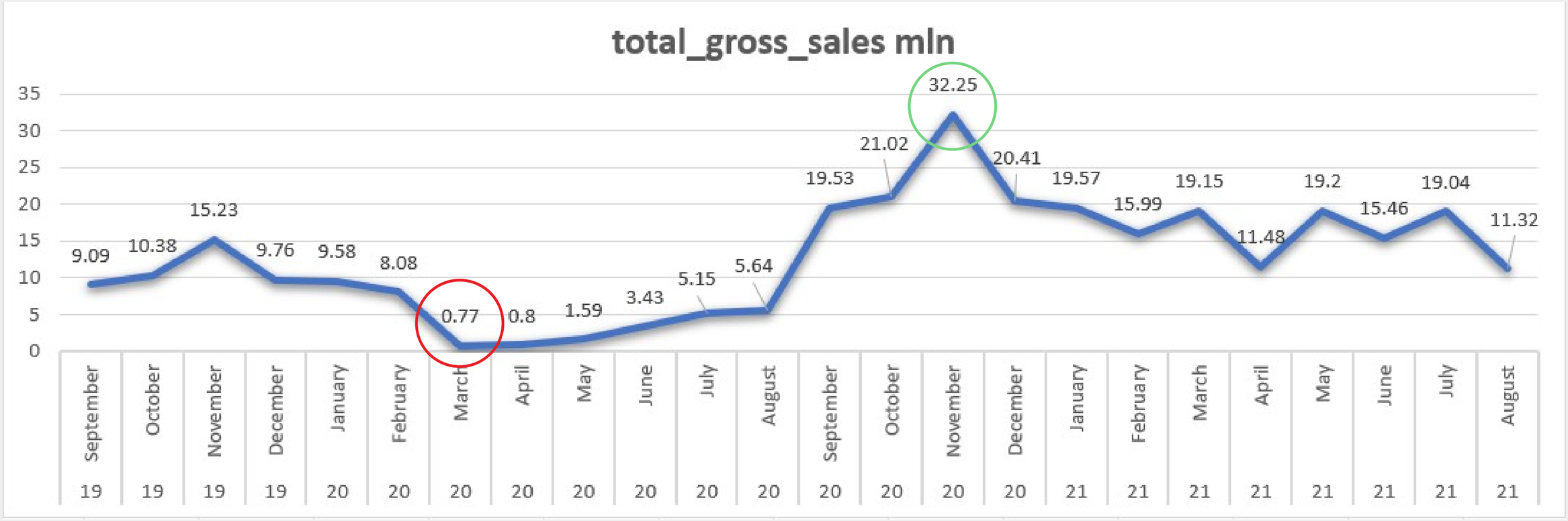
7. Get the complete report of the Gross sales amount for the customer “Atliq Exclusive” for each month. This analysis helps to get an idea of low and high-performing months and take strategic decisions. The final report contains these columns: Month,Year,Gross sales Amount

```sql
SELECT
    year(fm.date) as calender_year,
    CONCAT(MONTHNAME(fm.date), ' (', YEAR(fm.date), ')') AS 'month_m',
    round(SUM(fm.sold_quantity * fg.gross_price),2) AS total_gross_sales
FROM gdb023.fact_sales_monthly fm
JOIN fact_gross_price fg ON fm.product_code = fg.product_code
join dim_customer ds on
ds.customer_code=fm.customer_code
where ds.customer="Atliq Exclusive"
GROUP BY
    month_m,
    fm.date
order by fm.date,month_m;
```

calender_year	month_m	total_gross_sales
2019	September (2019)	9.09
2019	October (2019)	10.38
2019	November (2019)	15.23
2019	December (2019)	9.76
2020	January (2020)	9.58
2020	February (2020)	8.08
2020	March (2020)	0.77
2020	April (2020)	0.80
2020	May (2020)	1.59
2020	June (2020)	3.43
2020	July (2020)	5.15
2020	August (2020)	5.64
2020	September (2020)	19.53
2020	October (2020)	21.02
2020	November (2020)	32.25
2020	December (2020)	20.41
2021	January (2021)	19.57
2021	February (2021)	15.99
2021	March (2021)	19.15
2021	April (2021)	11.48
2021	May (2021)	19.20
2021	June (2021)	15.46
2021	July (2021)	19.04
2021	August (2021)	11.32

Insights:

Atliq Exclusive had the most sales in November 2020, and the lowest sales in March 2020. The sales dropped between March and August because of COVID-19. But the good news is that sales went up quickly after August and reached the highest level in November, which is the best in the past two years.



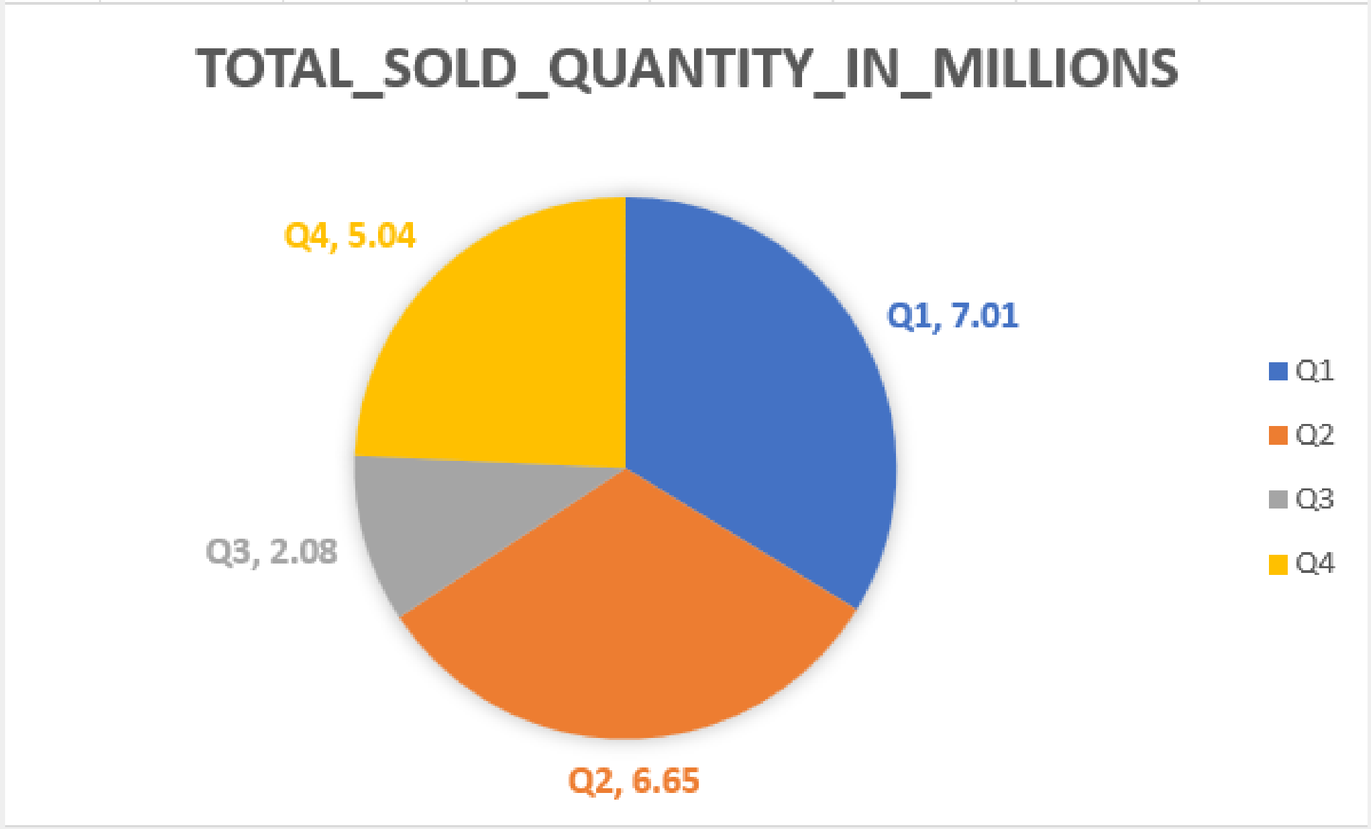
8. In which quarter of 2020, got the maximum total_sold_quantity? The final output contains these fields sorted by the total_sold_quantity, Quarter, total_sold_quantity

```sql
WITH cte1 AS (
    SELECT *,
        CASE
            WHEN MONTH(date) in (9,10,11) then CONCAT('Q1')
            WHEN MONTH(date)in (12,1,2) then CONCAT('Q2')
            WHEN MONTH(date) in (3,4,5) then  CONCAT('Q3')
            ELSE CONCAT('Q4')
        END AS fiscal_quarter
    FROM fact_sales_monthly
)

SELECT fiscal_quarter, SUM(sold_quantity) / 1000000 AS total_sold_quantity_in_millions
FROM cte1
WHERE fiscal_year = 2020
GROUP BY fiscal_quarter
ORDER BY total_sold_quantity_in_millions desc;
```

fiscal_quarter	total_sold_quantity_in_millions
Q1	7.0056
Q2	6.6496
Q4	5.0425
Q3	2.0751

Insights:

This connects with what we already talked about. It's about how COVID-19 affected our sales. In the third quarter of fiscal year 2020, which covers March, April, and May when COVID-19 was widespread, we sold fewer products, around 2.1 million. But we started getting better quickly even though the pandemic was still going on. This recovery in the fourth quarter might be because many students needed things like computers and notebooks for online classes, and there was a big demand for computer accessories too.



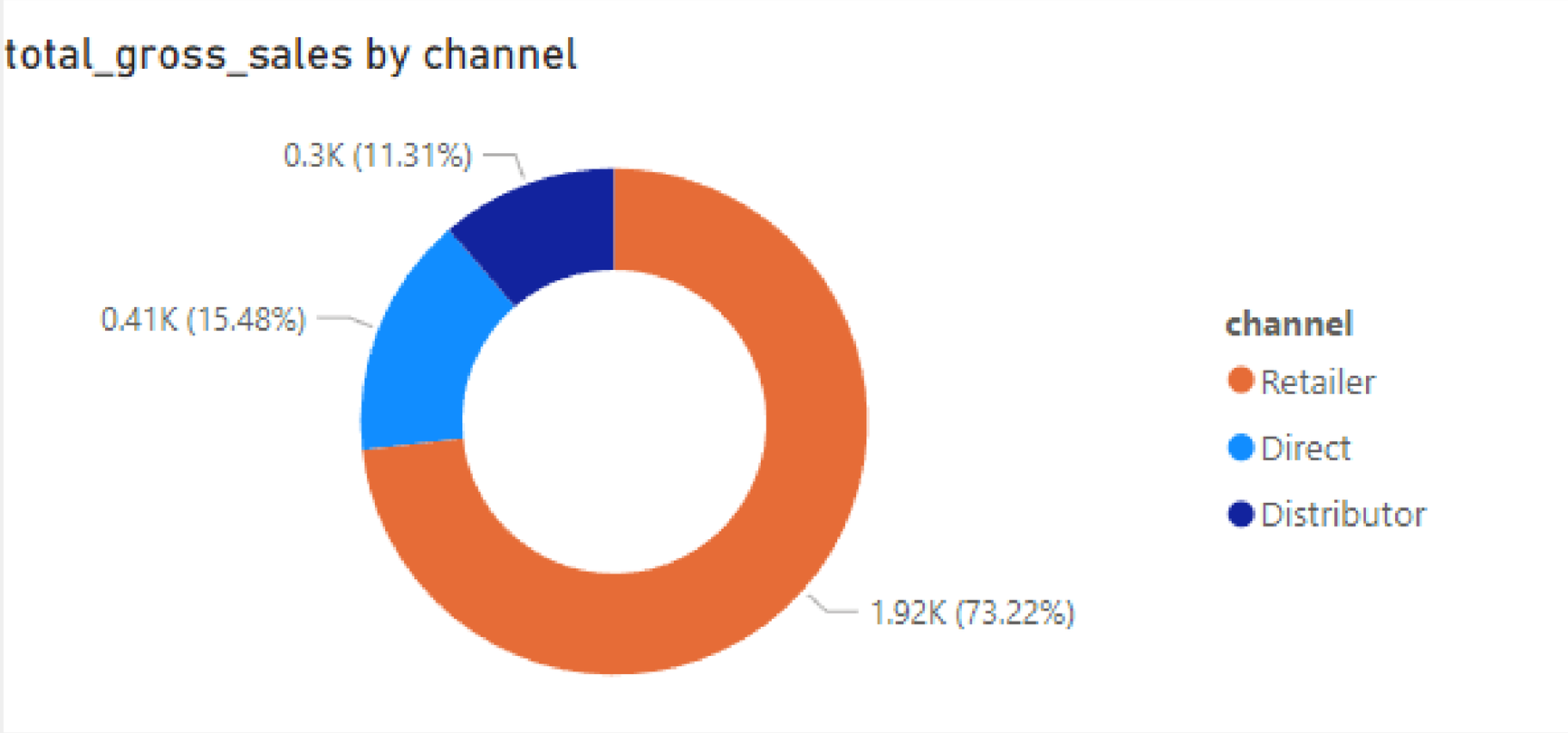
9. Which channel helped to bring more gross sales in the fiscal year 2021 and the percentage of contribution? The final output contains these fields, channel gross_sales_mln, percentage

```sql
with cte1 as (
SELECT
    ds.channel,
    round(SUM(fm.sold_quantity * fg.gross_price)/1000000,2) AS total_gross_sales
FROM gdb023.fact_sales_monthly fm
JOIN fact_gross_price fg ON fm.product_code = fg.product_code
join dim_customer ds on
ds.customer_code=fm.customer_code
where fm.fiscal_year=2021
GROUP BY
    ds.channel)

select channel,total_gross_sales,total_gross_sales*100/sum(total_gross_sales) over() as percentage_contro
from cte1
group by channel;
```

channel	total_gross_sales	percentage_contro
Direct	406.69	15.475031
Distributor	297.18	11.308047
Retailer	1924.17	73.216922

Insights:

"Retailer," brought in the most sales for the company, making up 73.22% of all sales. On the other hand, "Distributor" was the way that brought in the least sales, making up only 11.31% of the total.



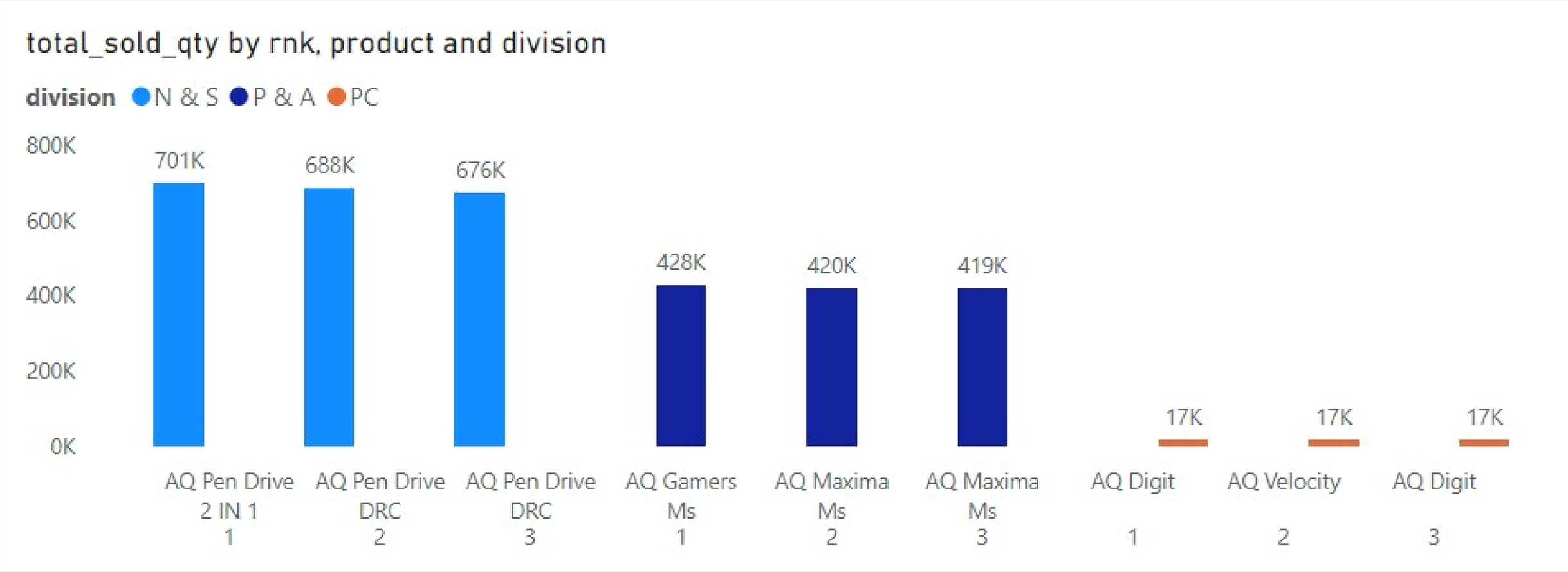
10. Get the Top 3 products in each division that have a high total_sold_quantity in the fiscal_year 2021? The final output contains these fields, division product_code,product ,total_sold_quantity ,rank_order

```sql
WITH cte1 AS (
    SELECT
        dp.division,
        dp.product_code,
        dp.product,
        SUM(fs.sold_quantity) AS total_sold_qty
    FROM gdb023.dim_product dp
    JOIN fact_sales_monthly fs ON dp.product_code = fs.product_code
    WHERE fiscal_year = 2021
    GROUP BY dp.division, dp.product_code, dp.product
),
ranked AS (
    SELECT
        division,
        product_code,
        product,
        total_sold_qty,
        DENSE_RANK() OVER (PARTITION BY division ORDER BY total_sold_qty DESC) AS rnk
    FROM cte1
)

SELECT division, product_code, product, total_sold_qty, rnk
FROM ranked
WHERE rnk <= 3;
```

division	product_code	product	total_sold_qty	rnk
N & S	A6720160103	AQ Pen Drive 2 IN 1	701373	1
N & S	A6818160202	AQ Pen Drive DRC	688003	2
N & S	A6819160203	AQ Pen Drive DRC	676245	3
P & A	A2319150302	AQ Gamers Ms	428498	1
P & A	A2520150501	AQ Maxima Ms	419865	2
P & A	A2520150504	AQ Maxima Ms	419471	3
PC	A4218110202	AQ Digit	17434	1
PC	A4319110306	AQ Velocity	17280	2
PC	A4218110208	AQ Digit	17275	3

Insights:

The most popular products in N&S were pen drives, with about 700,000 sold. In P&A, the top products were mice, with around 400,000 sold. For PC, the top products were personal laptops, and about 17,000 of them were sold.



Thank You